Original Study

Multimodal Assessment of Estrogen Receptor mRNA Profiles to Quantify Estrogen Pathway Activity in Breast Tumors

Anita Muthukaruppan,[1] Annette Lasham,[2] Kathryn J. Woad,[1] Michael A. Black,[3] Cherie Blenkiron,[2] Lance D. Miller,[4] Gavin Harris,[5] Nicole McCarthy,[6] Michael P. Findlay,[6] Andrew N. Shelling,[1] Cristin G. Print[2,7,8]

Abstract

We investigated the transcriptional pathways activated by estrogen receptor-α and their relationship with patient survival after treatment with endocrine therapy. From this analysis, we propose that multimodal assessment of breast tumors, using a combination of estrogen receptor-α status, estrogen receptor-α mRNA expression, and genomic indicators of estrogen pathway activity, could be useful for both research and treatment stratification.

Background: Molecular markers have transformed our understanding of the heterogeneity of breast cancer and have allowed the identification of genomic profiles of estrogen receptor (ER)-α signaling. However, our understanding of the transcriptional profiles of ER signaling remains inadequate. Therefore, we sought to identify the genomic indicators of ER pathway activity that could supplement traditional immunohistochemical (IHC) assessments of ER status to better understand ER signaling in the breast tumors of individual patients. **Materials and Methods:** We reduced *ESR1* (gene encoding the ER-α protein) mRNA levels using small interfering RNA in ER^+ MCF7 breast cancer cells and assayed for transcriptional changes using Affymetrix HG U133 Plus 2.0 arrays. We also compared 1034 ER^+ and ER^- breast tumors from publicly available microarray data. The principal components of ER activity generated from these analyses and from other published estrogen signatures were compared with ESR1 expression, ER-α IHC, and patient survival. **Results:** Genes differentially expressed in both analyses were associated with ER-α IHC and ESR1 mRNA expression. They were also significantly enriched for estrogen-driven molecular pathways associated with *ESR1*, cyclin D1 (*CCND1*), *MYC* (v-myc avian myelocytomatosis viral oncogene homolog), and *NFKB* (nuclear factor kappa B). Despite their differing constituent genes, the principal components generated from these new analyses and from previously published ER-associated gene lists were all associated with each other and with the survival of patients with breast cancer treated with endocrine therapies. **Conclusion:** A biomarker of ER-α pathway activity, generated using ESR1-responsive mRNAs in MCF7 cells, when used alongside ER-α IHC and ESR1 mRNA expression, could provide a method for further stratification of patients and add insight into ER pathway activity in these patients.

Keywords: Breast cancer, ER, *ESR1*, Gene expression, MCF7, Principal component analysis, RNA

K.J. Woad is currently affiliated with School of Veterinary Medicine and Science, University of Nottingham, Sutton Bonington Campus, Loughborough, Leicestershire, UK.

N. McCarthy is currently affiliated with ICON Cancer Care Wesley, Auchenflower, QLD, Australia.

[1]Department of Obstetrics and Gynaecology
[2]Department of Molecular Medicine and Pathology, Faculty of Medical and Health Sciences, The University of Auckland, Auckland, New Zealand
[3]Department of Biochemistry, University of Otago, Dunedin, New Zealand
[4]Department of Cancer Biology, Wake Forest University School of Medicine, Winston-Salem, NC
[5]Canterbury Health Laboratories, Christchurch, New Zealand
[6]Discipline of Oncology, Faculty of Medical and Health Sciences
[7]New Zealand Bioinformatics Institute
[8]Maurice Wilkins Centre for Molecular Biodiscovery, The University of Auckland, Auckland, New Zealand

Submitted: Jul 19, 2016; Revised: Aug 25, 2016; Accepted: Sep 2, 2016

Address for correspondence: Anita Muthukaruppan, PhD, Department of Obstetrics and Gynaecology, Faculty of Medical and Health Sciences, The University of Auckland, Private Bag 92019, Auckland 1142, New Zealand
E-mail contact: a.muthukaruppan@auckland.ac.nz

http://dx.doi.org/10.1016/j.clbc.2016.09.001

Introduction

Breast cancer is the most common cancer among women and the leading cause of cancer death in women worldwide. It is a multifactorial disease, with considerable interpatient heterogeneity and complex etiology involving different genetic, endocrinologic, environmental, and lifestyle influences.[1] The hormone estrogen plays a pivotal role in the pathogenesis of breast cancer,[2] and the determination of estrogen receptor-α (ER-α) status of a breast tumor using immunohistochemistry (IHC) is a longstanding molecular diagnostic test used to stratify breast cancer patients for endocrine-based treatment. The identification of ER-α in breast tumors using IHC predicts an increased likelihood of a response to endocrine-based therapy. Endocrine therapy is given to > 95% of patients with ER^+ tumors; however, a proportion of these patients will not have a treatment response or the cancer will become resistant to tamoxifen or other endocrine therapies over time.[3]

Although ER IHC is the current standard of care to determine ER status, concerns have been raised regarding inconsistencies in assigning ER status using IHC. It has been estimated that ≤ 20% of IHC assessments of ER status might be inaccurate (false-positive or false-negative result) because of a number of factors such as variations caused by IHC analysis, including tissue handling, IHC fixation methods, antigen retrieval techniques, antibodies used and their affinity and specificity, scoring practices, positivity cutoff thresholds, and interpretation criteria[4,5] (also described in 4 reviews[6-9]). It is also possible that some ER^- tumors might express functional variant ER-α protein isoforms that are not detected during IHC.[10,11]

Hence, international guidelines were established by the American Society of Clinical Oncology/College of American Pathologists to standardize the laboratory protocols for ER, progesterone receptor (PGR), and Erb-b2 receptor tyrosine kinase 2/HER2 (ERBB2) testing in breast cancer. The threshold for ER staining positive and the determination of ER status was set at 1% positive staining,[4,9,12,13] because it has been shown that patients with tumors with as little as 1% ER positivity will benefit from endocrine therapy.[4,14,15] Despite these useful guidelines, room still exists for variability in the determination of ER/PGR status of tumors. For example, the guidelines recommend the use of a number of antibodies such as the commonly used 1D5, 6F11, and the more sensitive SP1 for ER, and 3 different antibodies for PGR.[16] Furthermore, tumors can be scored using multiple methods, including the H-score, Allred score, and a quick score.[16] Although IHC is currently the reference standard method used to ascertain ER status in breast cancer, IHC provides no assessment of whether the estrogen signaling pathway targeted by endocrine therapy is actually functional in individual tumors. Therefore, a number of studies are now attempting to address this issue using gene expression profiling to more accurately define ER status.[14,17-22]

Gene expression profiling research using primary breast tumors[23,24] has transformed our understanding of the molecular heterogeneity of breast cancer. Many genomic studies now integrate analysis of primary cancers using various genomic, transcriptomic, proteomic, and clinical parameters, such as exome and microRNA sequencing, tumor mutational data, DNA copy number and methylation, mRNA expression and protein expression arrays, and in vitro experimental data, to investigate various aspects of cancer.[25-28] In breast cancer, integrated studies have now enabled, for example, the classification of breast cancer using RNA and DNA profiling,[29] the investigation of pathway signaling in breast cancer,[30] the association of DNA number variations with gene expression and patient survival,[31] and the investigation of in vivo biology from in vitro experiments.[32,33] Genomic profiling studies have also facilitated breast cancer prognostication and stratified therapy,[24,34] including first-generation signatures, such as MammaPrint (Agendia BV, Amsterdam, Netherlands), Onco*type* DX (Genomic Health, Inc, Redwood City, CA), and the Genomic Grade Index (MapQuant Dx, Ipsogen, France) to the more recent Prosigna Assay based on the PAM50 test (NanoString Technologies, Inc, Seattle, WA), Breast Cancer Index (BioTheranostics, San Diego, CA), and Endopredict (Sividon Diagnostics GmbH, Koln, Germany). However, evidence is still being gathered regarding the clinical utility of breast cancer genomic assays[35-38]; therefore, traditional nongenomic pathology examination remains a key component of the standard of care.

In general, good concordance has been reported between the ER status determined by IHC and the ER status ascertained using MapQuant,[39] TargetPrint,[40] and 3 molecular assays, BluePrint, MammaPrint, and TargetPrint.[41] Allott et al[42] also found high concordance for ER status of >10% positivity between IHC and RNA-based intrinsic subtypes.[42] However, in these studies, the concordance rates for PGR status were always lower than those for ER status. Possible implications exist when deciding on the treatment options for patients when the hormone receptor status determined using IHC or molecular tests are not in accordance. Thus, the use of microarray data to assign ER/PGR status still needs to be refined.

ESR1 (the *ESR1* gene encodes the ER-α protein) mRNA expression drives global patterns of gene expression in breast cancer cells.[43] Despite years of investigation, our understanding of the ER signaling pathways that mediate these gene expression programs is still incomplete. Breast tumors are usually designated as either ER^+ or ER^- according to the percentage of cells with detectable ER-α staining (and, indirectly, the intensity of the staining within the positive cells). In contrast, ESR1 mRNA expression is a continuous variable when measured using microarray or reverse transcription-quantitative polymerase chain reaction (RT-qPCR) and exhibits a wide range of values.[18,19,44] This range of ESR1 mRNA expression could likely translate into a range of ER transcriptional activity in a breast tumor and might, along with intratumor ER pathway heterogeneity, influence a patient's response to endocrine therapy.[45] For example, low expression of ESR1 mRNA in ER^+ breast cancer has been associated with tamoxifen resistance,[46] with high levels of ESR1 mRNA expression identified by in situ hybridization associated with tamoxifen responsiveness.[21] The activity of transcription factors such as the ER can be inferred by measuring the expression levels of sets of their transcriptional target genes,[47] which are often statistically summarized into a single value reflecting transcription factor activity, such as a "principal component" (PC).[48] The concept of using statistical summaries of the expression of gene sets, including PC analysis, to estimate the activity of specific molecular pathways has been widely used to study breast cancer, such as triple-negative breast cancer.[49] This type of analysis is also useful when attempting to integrate data from multiple

sources, such as DNA methylation and gene expression data[50] or studying the relationship between single nucleotide polymorphisms and trastuzumab resistance in breast cancer cell lines.[51] Several gene sets associated with ER signaling have been identified for use in analysis. These include gene sets that can differentiate between ER^+ and ER^- tumors,[52] estrogen-regulated genes that predict survival or recurrence of patients with ER^+ cancer treated with tamoxifen,[46,53-56] and the estrogen receptor attractor metagenes from the 2012 DREAM (Dialogue for Reverse Engineering Assessments and Methods) Breast Cancer Prognosis Challenge.[57]

It is becoming evident that the assessment and determination of ESR1 mRNA expression levels in a tumor are also important, in addition to the IHC-determined ER status (based on ER-α protein expression). Moreover, an assessment of an intact estrogen signaling pathway is imperative to predict a patient's response to endocrine therapy. Despite numerous studies of ER-α signaling in breast cancer, the relationship between ER-α–driven transcriptional programs and its relationship with ER status and ESR1 mRNA transcript abundance and its distribution across a full range of breast tumors remains unclear.

Therefore, in the present study, we initially performed experiments to identify ER-α–dependent gene expression patterns in a simple and well-controlled cell culture model. Specifically, we mapped the gene expression changes after small interfering RNA (siRNA)-mediated reduction of ESR1 levels in ER^+ MCF7 breast cancer cells. Next, the expression of this gene set and 12 previously published ER-associated gene sets were interpreted in the context of 1034 patients with breast cancer. The distribution of these gene sets, ESR1 mRNA expression, and ER status (from IHC) were interpreted with other clinicopathologic data across breast tumor subtypes. From our analysis, we suggest that genomic indicators that summarize estrogen pathway activity from gene expression data, when used with IHC-determined ER status and ESR1 mRNA expression, can provide a useful profile of ER signaling in the breast tumors of individual patients.

Materials and Methods

Cell Culture and Transfection of MCF7 Cells With siRNAs

MCF7 breast cancer cells were purchased from the American Type Culture Collection (Manassas, VA) and cultured in Dulbecco's modified Eagle medium (Thermo Fisher Scientific Inc, Waltham, MA), 10% vol/vol fetal bovine serum (Thermo Fisher Scientific Inc), 0.01 mg/mL bovine insulin (Sigma-Aldrich, St Louis, MO), 2 mM L-glutamine (Thermo Fisher Scientific Inc), 1 mM sodium pyruvate (Thermo Fisher Scientific Inc), 1 U/mL penicillin G (Thermo Fisher Scientific Inc), and 1 μg/mL streptomycin sulfate (Thermo Fisher Scientific Inc) in humidified air with 5% vol/vol carbon dioxide at 37°C.

A new batch of MCF7 cells was obtained from the American Type Culture Collection for the present study and used within 10 passages of purchase. Stealth-modified siRNA duplex HSS176619 (Thermo Fisher Scientific Inc) was used to target *ESR1*; BLOCK-iT Fluorescent Oligo (Thermo Fisher Scientific Inc) and Stealth RNAi Negative Control Medium GC Duplexes (Thermo Fisher Scientific Inc) were used as positive and negative nontargeting transfection controls, respectively. Transfections were performed in triplicate using a "reverse transfection" method. A transfection mix containing 2.5 μL of RNAiMAX transfection reagent (Thermo Fisher Scientific Inc) and 5 nM siRNA in 500 μL of OptiMEM serum-free media (Thermo Fisher Scientific Inc) was incubated for 20 minutes at room temperature and added dropwise to 5.0×10^5 cells in 2.5 mL of antibiotic-free, complete growth media per well in a 6-well plate. The cells were incubated at 37°C for 24 hours, the media removed, and the cells harvested using 1 mL of TRIzol reagent (Thermo Fisher Scientific Inc). The details of the optimization procedure are provided in the Supplemental Methods (available in the online version).

RNA Extraction and Microarray Analysis of MCF7 Cells

Total RNA from MCF7 cells was isolated from TRIzol homogenates using chloroform and purified using the PureLink RNA mini kit according to the manufacturer's protocol (Thermo Fisher Scientific Inc). RNA yields and purity were determined using the NanoDrop spectrophotometer (NanoDrop Technologies Inc, Wilmington, DE), and RNA integrity was assessed using the Experion Automated Electrophoresis System according to the manufacturer's protocol (Bio-Rad Laboratories, Berkeley, CA). The average 260/280 purity ratio was 2.0 (range, 1.9-2.0), and the RNA quality indicator was 9.7 (range, 9.5-9.8). Total RNA (1 μg) was labeled, fragmented, and hybridized to Affymetrix Human Genome U133 Plus 2.0 arrays in accordance with the manufacturer's protocol (Affymetrix, Santa Clara, CA). Raw and normalized microarray data are available in the Gene Expression Omnibus (GEO) record GSE37820. Quality control was performed using Expression Console (Affymetrix) and dChip before quantile normalization using R statistical software (available at: http://cran.r-project.org/) and the robust multichip averaging (RMA) algorithm with background correction and loess splining, as implemented in the R "affy" package. The extent of *ESR1* mRNA knockdown identified using microarrays was confirmed using RT-qPCR (see Supplemental Methods available in the online version).

Bioinformatic Analyses of Microarray and RNA Sequencing Data

Microarray data for 1034 primary breast tumors were assembled from raw Affymetrix Human Genome U133 ".cel" files, available on GEO: GSE1456, GSE3494, GSE4922, GSE6532, GSE7390, GSE36771, GSE36772, and GSE36773. A total of 22,277 probe sets common to both the U133 Plus 2.0 and the U133A version 2 arrays were used for the present analysis, as described previously.[58] Array quality assessment was performed using the "AffyQCReport" package in R and normalized using RMA. Normalized gene level RNA sequencing (RNAseq) and clinical data from 1023 breast tumors were downloaded from the Cancer Genome Atlas data portal (available at: https://tcga-data.nci.nih.gov/tcga/) and processed using the "rsem" method in R. Probe sets differentially expressed between experimental samples were ranked using the linear models for microarray data (LIMMA) package in R. The lists of differentially expressed probe sets were tested for enrichment of particular functional categories using ingenuity pathways analysis (Ingenuity Systems; available at: http://www.ingenuity.com). The ESR1 probe set 205225_at was used as the indicator of ESR1 expression in all analyses of microarray data (see Supplemental Methods available in

the online version). The percentile rank of ESR1 mRNA expression was calculated from the $\log_2$ expression data by determining the relative expression level of ESR1 probe set 205225_at, ranked against all other probe sets on the array.

Gene expression data from ESR1 siRNA-treated MCF7 cells were used to develop a putative biomarker of estrogen activity in clinical samples. The genes significantly regulated in MCF7 cells by ESR1 siRNA treatment were identified. Subsequently, the data corresponding to these genes were extracted from clinical samples and a PC was inferred from these data. The PC analysis was performed by singular value decomposition using the "svd" function in R, after the data had been zero centered and scaled to unit variance. Because the signs of PCs can be inconsistent, in all cases, PC signs were set such that the PC correlation with ESR1 expression was > 0. Statistical meta-analysis was performed using the "metaMA" package in R. Molecular subtypes were assigned to each tumor using the single sample predictor algorithm applied per cohort.[59] Spearman's correlation coefficients between any 2 probe sets were calculated using the "cor" function in R. The intersection of genes between selected numbers of gene sets were plotted using our own modifications of the "heatmap.2" function of the "gplots" package in R. Kaplan-Meier survival analysis with log-rank tests and Cox proportional hazards analysis was performed using the "survival" package in R (see Supplemental Methods for a detailed description available in the online version). The visualizations for many of the analyses presented in the present report were scripted using the R "shiny" package from the R studio framework (available at: http://www.rstudio.com/products/shiny/shiny-server/).

Results

Effect of ESR1 siRNA Treatment on ESR1 Expression in MCF7 Cells

To identify the gene expression patterns driven by ER-α in a controlled in vitro system, we treated cultures of MCF7 cells with an siRNA to *ESR1*. The average knockdown of ESR1 was 72.5% and 78.6% as assessed by microarray analysis (Figure 1A) and RT-qPCR (Figure 1B), respectively. Before and after ESR1 mRNA knockdown in MCF7 cells, the mean percentile ranks (see Materials and Methods) of ESR1 expression on the microarrays were 93% and 76%, respectively. These percentile ranks were then mapped to the expression percentile ranks of the same Affymetrix ESR1 probe set 205225_at in 1034 breast tumors, separated into ER^- and ER^+ (Figure 1C). Before ESR1 knockdown, ESR1 expression in MCF7 cells was at a level typical of most ER^+ luminal tumors. However, when the analysis was restricted to only luminal A tumors (Figure 1D), it was evident that after siRNA treatment, ESR1 expression was reduced to a level less than that of any ER^+ luminal A tumors. This suggests that ESR1 siRNA treatment of MCF7 cells allows a comparison between cells with ESR1 levels typical of most ER^+ luminal A tumors and cells with ESR1 levels less than those found in any ER^+ luminal A tumor.

Genes Showing Altered Expression After ESR1 siRNA Treatment of MCF7 Cells

The experimental objective was to identify differential mRNA expression between MCF7 cells in which ER-α–driven gene expression responses were high versus low. Therefore, to allow ER-α–mediated gene expression programs to be driven by steroid hormone components of culture media and by the endogenous estrogens produced by the cells, we followed a procedure similar to that used by others,[60,61] in which no steroid hormone depletion steps were performed (ie, serum was not charcoal stripped and phenol red was not removed from the media), and no exogenous estrogen was added to the media. Analysis of the gene expression data in MCF7 cells after ESR1 knockdown using LIMMA identified 50 Affymetrix probe sets that were differentially expressed (multiple testing-adjusted threshold set at $P \leq .001$; Supplemental Table 1; available in the online version). These mRNAs consisted of 39 unique annotated genes (some could not be annotated or were duplicate probe sets) that included *ESR1* targets such as the *PGR* anterior gradient 2 homolog (*AGR2*), carbonic anhydrase XII (*CA12*) and *MYC* (v-myc avian myelocytomatosis viral oncogene homolog). These genes are referred to as the MCF7-*ESR1*-siRNA gene set. This gene set appears to constitute several interconnected molecular pathways involving *ESR1*, *CCND1* (cyclin D1), *MYC*, and *NFKB* (nuclear factor kappa B). The multiple testing-adjusted enrichment threshold was set at $P \leq .05$. The genes constituting these canonical pathways are listed in Supplemental Table 2 (available in the online version).

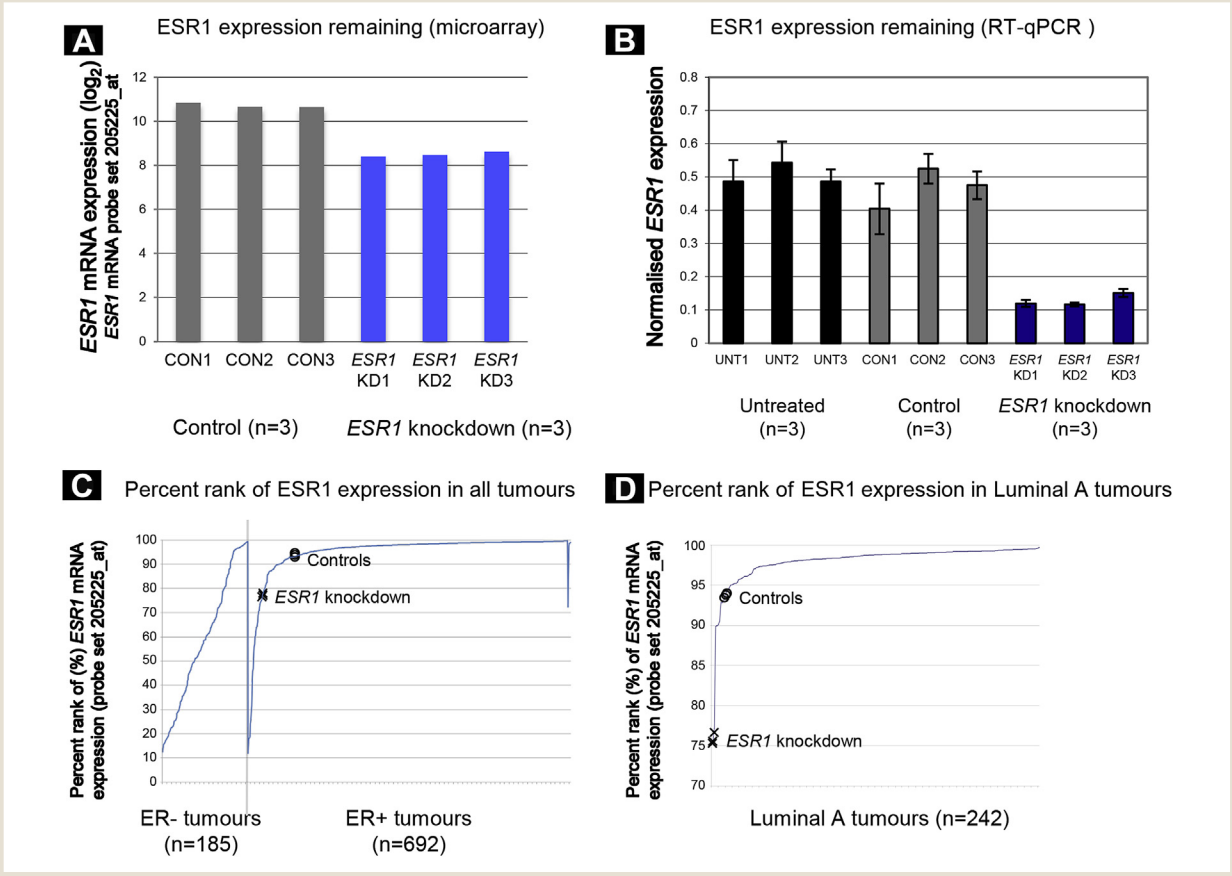
*Generation of PC of ER-α Pathway Activity From the MCF7-*ESR1*-siRNA Gene Set*

The expression of genes constituting the MCF7-*ESR1*-siRNA gene set was examined in microarray data from 1034 breast tumors assembled from several published cohorts (see Materials and Methods). The expression values of 28 of the 39 members of this gene set, for which data could be generated across all 1034 breast tumors, are shown as a heatmap in Figure 2. A first PC (excluding ESR1 and including PGR; see Materials and Methods) was generated to provide a statistical summary of these mRNAs for each tumor and was used to sort the tumors across the heatmap in Figure 2. We hypothesized that this PC of the MCF7-ESR1-siRNA gene set might provide a robust indicator or biomarker of ER-α transcriptional activity in breast tumors. Sorting tumors according to the magnitude of this first PC also appeared to segregate the tumors, in general, by molecular subtype, grade, and ER status (Figure 2). Tumors in the lowest 25% of PC magnitude were mostly grade 2 or 3 tumors and of the basal-like or ERBB2 subtypes, with 87% being ER^- by IHC. In contrast, the tumors in the highest 75% of PC magnitude were mostly luminal A, luminal B, or normal-like subtypes, with 95% ER^+.

We also noted that although most ER^+ tumors had greater ESR1 mRNA expression levels than ER^- tumors, the ESR1 mRNA expression range of ER^+ and ER^- tumors overlapped considerably (ESR1 microarray probe set 205225_at; Supplemental Figure 1; available in the online version). Similarly, an ESR1 mRNA expression overlap was present among the breast cancer molecular subtypes, with the distributions most similar between luminal A and luminal B tumors (Supplemental Figure 2; available in the online version).

The MCF7-ESR1-siRNA first PC for 5 of the individual microarray study cohorts (Figure 3A-E) and as a combined data set of all 5 cohorts (n = 1034; Figure 3F) correlated with ESR1 mRNA abundance. In addition to consistency across independent

Figure 1 Assessment of ESR1 (Gene Encoding Estrogen Receptor [ER]-α Protein) mRNA Expression After Small Interfering RNA (siRNA) Knockdown in MCF7 Cells. (A) Log_2 Expression Levels of ESR1 mRNA Measured by Microarray From Hybridization to Probe Set 205225_at. Each Biologic Sample Was Assayed in Triplicate. (B) Normalized Measurement of ESR1 mRNA Levels in Untransfected, Control siRNA- or ESR1 siRNA-Transfected MCF7 Cells Using Reverse Transcription Quantitative Polymerase Chain Reaction. (C) Log_2 ESR1 mRNA Expression Levels Shown as Percent Rank (Percentile of Expression) in MCF7 Cells Transfected With siRNA Control (Circles) or ESR1 siRNA (Crosses) Were Plotted With All Breast Tumors (Sorted by Increasing Percent Rank of ESR1 Expression) and Shown Separately as ER^- and ER^+ Tumors (by Immunohistochemistry), and (D) Luminal A Breast Tumors (Sorted by Increasing Percent Rank of ESR1 Expression)



Abbreviations: CON = control; KD = knockdown; UNT = untreated.

microarray data sets, this relationship was also consistent across analysis platforms (microarray vs. RNAseq). The MCF7-ESR1-siRNA PC correlated with the ESR1 mRNA abundance from 1023 breast tumors from the Cancer Genome Atlas breast cancer RNAseq data set (Supplemental Figure 3 in the online version). Both Figure 3F and Supplemental Figure 3 (available in the online version) revealed that a small proportion of patients assessed by IHC to be ER^- actually had high ESR1 mRNA expression and high activity of the ER pathway as inferred by PC analysis. In contrast, a small proportion of patients determined to be ER^+ by IHC expressed low levels of ESR1 mRNA and low inferred activity of the ER pathway. These 2 patient groups are potentially clinically relevant (see Discussion). Analysis of the ESR1 mRNA abundance and ER IHC status plotted with the MCF7-ESR1-siRNA PC distributed across the 5 breast cancer subtypes in the 1034 breast tumors showed that the ESR1 mRNA expression and the MCF7-ESR1-siRNA PC progressively increased from basal-like, ERBB2, and normal-like to luminal tumors (Figure 4). The MCF7-ESR1-siRNA PC also correlated well with ESR1 mRNA and IHC PGR status (Figure 5A) and PGR mRNA and IHC PGR status (Figure 5B) across all breast tumors. A wide range of ER pathway activity was found in IHC PGR^- tumors that also had very low PGR mRNA expression (Figure 5B).

Comparison of Genes Differentially Expressed After ESR1 *Knockdown and Other Estrogen Pathway-Associated Gene Sets With Patient Survival*

We then identified 23 Affymetrix probe sets, each corresponding to a unique gene, that were differentially expressed between ER^+ and ER^- tumors from the collated breast cancer cohort (multiple testing corrected, $P = .001$; genes listed in Supplemental Table 3; available in the online version). We assessed the intersection between this gene set with the MCF7-ESR1 gene set and 12 other previously published gene sets

Figure 2 Heatmap Depicting mRNA Expression of Probe Sets of MCF7-*ESR1* (Gene Encoding Estrogen Receptor [ER]-α Protein)-siRNA (Small Interfering RNA) Gene Set. Expression Values of 28 of 33 Members of MCF7-ESR1-siRNA Gene Set for Which Data Could Be Generated Across All 1034 Breast Tumors Were Used to Generate a Principal Component (PC). Where Multiple Probe Sets Were Available for Any 1 Gene, These Were Averaged. Tumors (n = 1034) Were Arranged From Low PC Magnitude (Green) to High PC Magnitude (Red; Left to Right). Genes Used to Infer the PC of ER Signaling Were Hierarchically Clustered Using Euclidean Distance and Ward's Method Linkage (Dendrogram Left of Heatmap). Their mRNA Expression Levels Across the Breast Tumors Are Indicated in the Heatmap (High, Red; Low, Green). The Centroid Refers to the Mean Expression Within Each Tumor

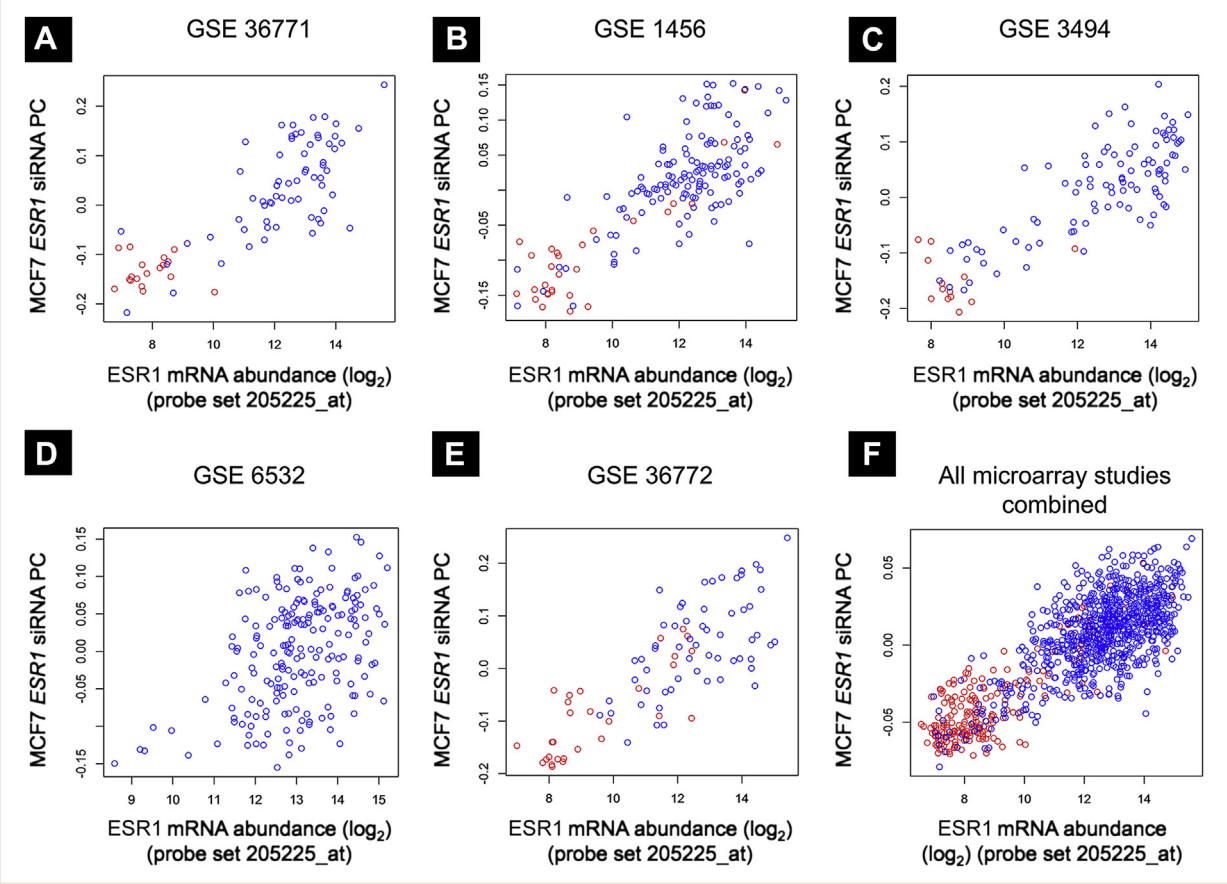
Abbreviations: ERBB2 = Erb-b2 receptor tyrosine kinase 2/HER2; IHC = immunohistochemistry.

representing different facets of the estrogen pathway in breast cancer (gene lists provided in Supplemental Table 3; available in the online version). The intersection of genes between gene sets that represented the estrogen response and purportedly related to estrogen function in breast cancer revealed little overlap with each other (Supplemental Figure 4; available in the online version). For example, the set of 39 genes differentially expressed after ESR1 knockdown in MCF7 cells and the set of 34 genes differentially expressed between ER^+ and ER^- tumors had only 1 gene in common, PGR.

To assess which aspects of ER biology are most highly associated with patient prognosis, we analyzed the distant metastasis-free survival (DMFS) of ER^+ patients from the 1034 breast cancer cohort, using the gene sets listed in Supplemental Table 3 (available in the online version). Analyses of 326 breast cancer patients who underwent endocrine therapy but not chemotherapy showed that the MCF7-ESR1-siRNA PC was weakly to moderately associated with DMFS in these patients (Cox proportional hazards, $P = .0001$); however, the association between the MCF7-ESR1-siRNA PC and DMFS was not statistically significant in untreated patients (Figure 6). Similarly, the PC of genes identified to be differentially expressed between ER^- and ER^+ tumors and PCs of the Endopredict,[55] PAM50,[62] and simplified tamoxifen predictor[56] gene sets were significantly associated with DMFS in endocrine therapy-treated patients but had a weaker association with DMFS in untreated patients (Figure 6; Supplemental Table 4; available in the online version). The ESR1 mRNA level itself and the PCs of the additional ER-associated gene sets showed a range of association with patient prognosis, in some cases showing a trend toward a stronger association after endocrine treatment than in untreated patients (Supplemental Table 4; available in the online version).

In the 1034 breast tumor data set, a group of ER^- IHC patients who had not received endocrine therapy nevertheless had high ESR1 mRNA expression and high inferred ER pathway activity (Figure 7A). Given their apparent ER pathway activation, it is possible that a subset of these patients might, in fact, have benefited from endocrine therapy. In contrast, a group of ER^+ patients who had received endocrine therapy had low ESR1 mRNA expression and low inferred ER pathway activity, possibly representing heterogeneous tumors with focal ER activation (Figure 7B).

Figure 3 Comparison of ESR1 (Gene Encoding Estrogen Receptor [ER]-α Protein) mRNA Expression, MCF7-ESR1-siRNA (Small Interfering RNA) First Principal Component (PC) and ER Status. Log_2 ESR1 mRNA Levels Were Plotted Against MCF7-ESR1-siRNA PC in Breast Tumors From 5 Selected Microarray Data Sets. (A) GSE36771 (n = 107), (B) GSE1456 (n = 159), (C) GSE3494 (n = 232), (D) GSE6532 (n = 241), (E) GSE36772 (n = 100), (F) All 5 Data Sets (n = 1034). ER Status Determined by Immunohistochemistry Shown as Red for ER^- Tumors and Blue for ER^+ Tumors



We then assessed whether these ER-associated gene sets might be able to provide additional prognostic or predictive information to ER IHC status or ESR1 mRNA information for breast cancer patients who had received endocrine therapy but not chemotherapy (see Supplemental Methods in the online version). The addition of first PCs from the following gene sets, to either ER IHC status or ESR1 mRNA levels, generated logistic regression models that might be superior to those based on either ER IHC status or ESR1 mRNA levels alone: MCF7-ESR1-siRNA first PC; the simplified tamoxifen predictor; or Endopredict (Supplemental Table 5; available in the online version). The addition of the first PC of all 3 of these gene sets increased the separation of the hazard ratio's 95% confidence intervals away from 1.0, reduced the model Akaike information criterion, and increased model C indexes for endocrine-treated patients. However, the effect of adding these gene set PCs to models based on ER IHC status or to ESR1 mRNA information for untreated patients was inconsistent, as was the effect of adding other ER-associated gene sets to ER status or ESR1 mRNA information for either endocrine-treated patients or untreated patients (Supplemental Table 5; available in the online version).

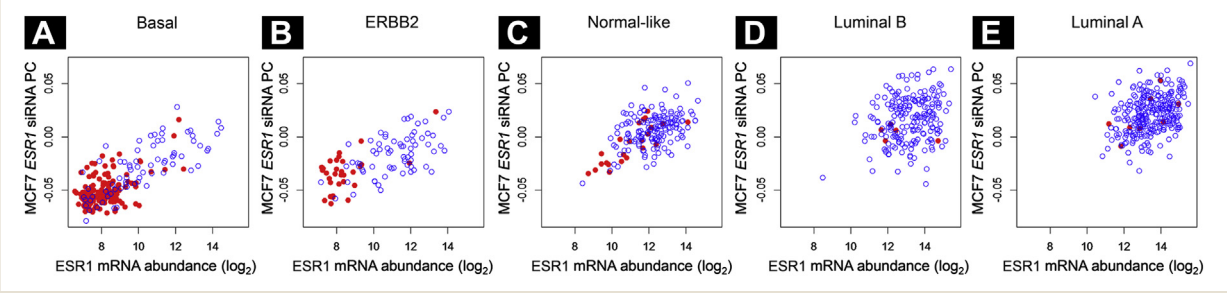
Discussion

ER-α–Regulated Molecular Pathways in MCF7 Cells

One method of investigating the underlying biology in cancer cell lines on a larger scale is to create as many different states of these cells as possible using an experimental perturbation approach and then to use computational methods, such as PC analyses or gene networks, to model the gene expression data from these different cell states.[63,64] By treating the cell line as a system, the responses from the system such as pathways and molecular interactions in cell perturbations such as the overexpression of mutant genes, the knockdown of single or multiple genes, or treatment with different hormones or chemotherapy drugs, can be analyzed. MCF7 breast cancer cells have been extensively used for in vitro perturbation experiments for extrapolation in vivo.[63] Our approach of knocking down target gene expression, in this case ESR1, in the MCF7 cells was reciprocal to the approach used by Bild et al.[65] They identified gene expression signatures that were associated with the activation of oncogenic pathways using adenoviruses in quiescent human mammary epithelial cells.[65]

In our study, we first investigated mRNA expression changes after siRNA knockdown of ESR1 in MCF7 cells using microarrays

Figure 4 Comparison of ESR1 (Gene Encoding Estrogen Receptor [ER]-α Protein) mRNA Expression, MCF7-ESR1-siRNA (Small Interfering RNA) Principal Component (PC) and ER Status Across Breast Cancer Subtypes. Log_2 ESR1 mRNA Levels Were Plotted Against MCF7-ESR1-siRNA PC in 1034 Breast Tumors Separated by Subtype. (A) Basal-Like (n = 209), (B) Erb-b2 Receptor Tyrosine Kinase 2/HER2-Positive (n = 100), (C) Normal-Like (n = 176), (D) Luminal B (n = 210), (E) Luminal A (n = 291; Tumors That Could Not Be Classified Into a Subtype Were Excluded). ER Status Determined by Immunohistochemistry Shown as Red for ER^- Tumors and Blue for ER^+ Tumors



to identify *ESR1*-regulated genes, including genes expressed because of estrogen-independent ER signaling. Although recognizing that MCF7 cells are derived from an advanced metastatic tumor, the relatively high expression of ESR1 in MCF7 cells is in concordance with our understanding that MCF7 cells more closely represent luminal A breast tumors than other subtypes[66,67] and with the widespread use of MCF7 cells as a model of ER^+ breast cancer. The mRNAs that showed altered abundance between control samples and ESR1 knockdown samples in MCF7 cells are components of known estrogen-driven molecular pathways, including pathways centered on *CCND1*, *ESR1*, *MYC*, and *NFKB*. These estrogen-driven signaling pathways are known to be important for tumor biology and are targets of endocrine therapy. For example, CCND1 has been shown to drive proliferation in breast cancer,[68] *MYC* pathways are often activated in luminal B tumors,[69] and NFKB signaling is involved in resistance to endocrine-based therapy.[70]

Our MCF7 ESR1 siRNA experimental system does have limitations. For example, it does not address signaling through the ER-β receptor, which might allow responses to tamoxifen in ER^- patients.[71] In addition, it does not address differential ER-α binding due to redistribution of ER-α, FOXA1 (forkhead box A1),[72] and other cofactors such as the occurrences of ER-α and other splice variants. The experimental system used in our analysis, like similar systems used previously by others,[60,61] contains no steroid hormone depletion steps (the serum was not charcoal stripped and phenol red was not removed from the media) and no addition of exogenous estrogens. In this system, ER-α–mediated gene expression programs are driven by steroid hormone components of culture media and by the endogenous

Figure 5 Comparison of ESR1 (Gene Encoding Estrogen Receptor [ER]-α Protein) siRNA (Small Interfering RNA) First Principal Component (PC) with ESR1 and Progesterone Receptor (PGR) mRNA Expression and PGR Status. (A) Log_2 ESR1 mRNA Levels Were Plotted Against MCF7-ESR1-siRNA PC in 1034 Breast Tumors. PGR Status Determined by Immunohistochemistry Shown as Red for PGR^- Tumors and Blue for PGR^+ Tumors. (B) Log_2 PGR mRNA Levels Were Plotted Against MCF7-ESR1-siRNA PC in 639 Breast Tumors (Remaining Tumors Had No PGR Status Available). PGR Status Determined by Immunohistochemistry Shown as Red for PGR^- Tumors and Blue for PGR^+ Tumors

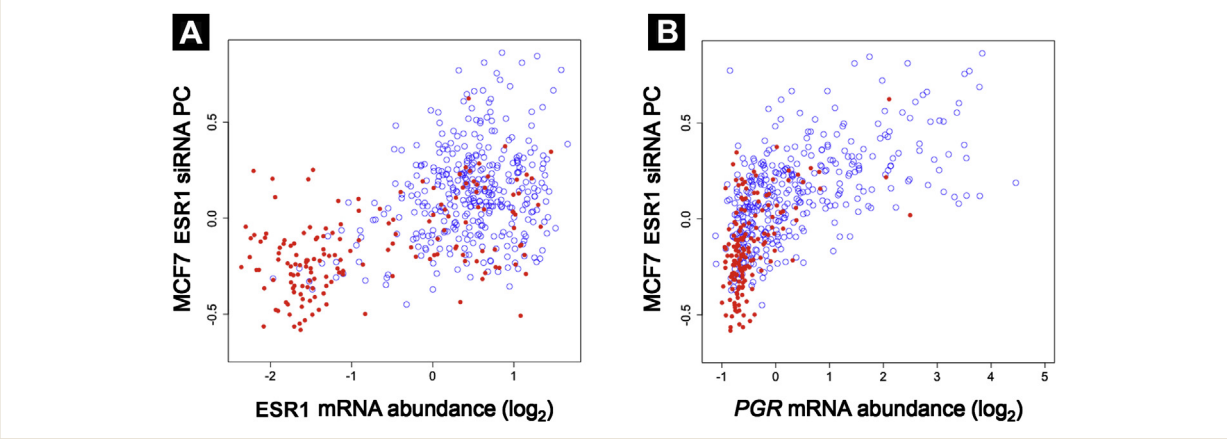


Figure 6 Association of Gene Sets With Distant Metastasis-Free Survival (DMFS) in Nonadjuvant Treated or Endocrine-Treated Patients. Kaplan-Meier Survival Plots Showing Association of Principal Component of 4 Gene Sets: MCF7-ESR1 (Gene Encoding Estrogen Receptor [ER]-α Protein)-siRNA (Small Interfering RNA), ER^+ Versus ER^-, Endopredict, PAM50, and Simplified Tamoxifen Predictor and Prognosis of Either Nonadjuvant-Treated or Endocrine-Treated Breast Cancer Patients. Cox Proportional Hazards (CoxPH) and Log-Rank Test (LR) *P* Values Shown for 4 Gene Sets That Showed the Most Significant Association With DMFS in Patients Receiving Endocrine Therapy. High Indicates Greater Than the Quantile Cutoff; and Low, Less Than the Quantile Cutoff

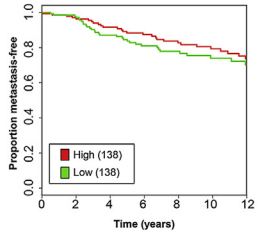
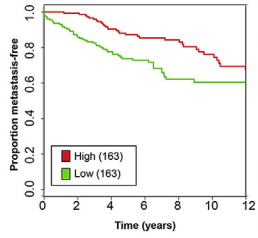
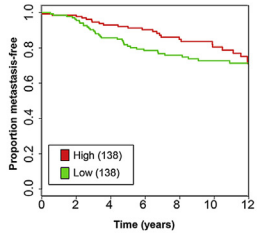
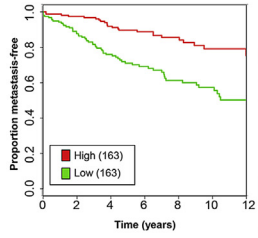
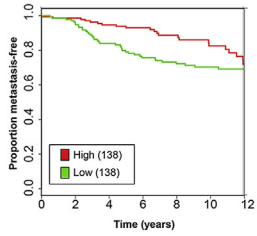
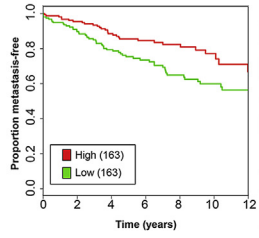
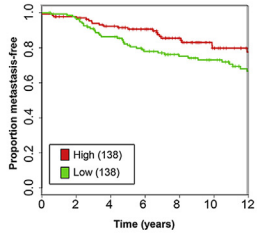
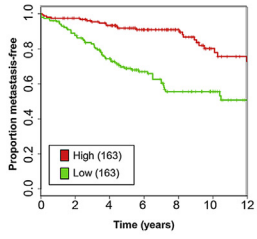
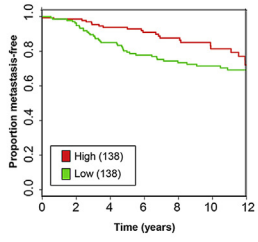
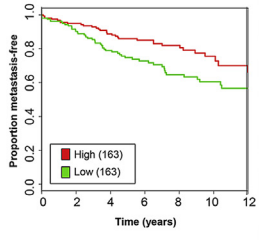
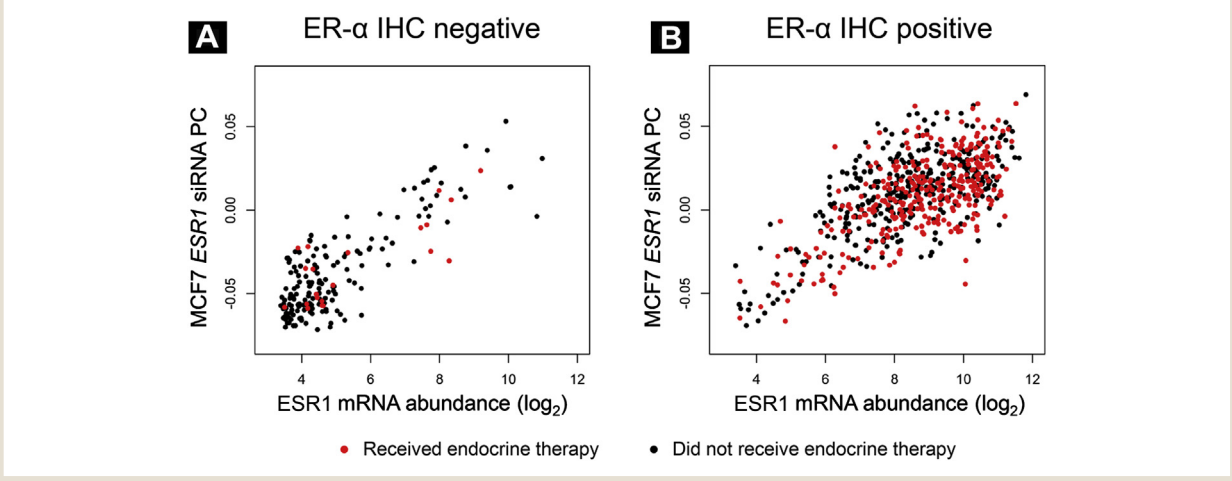
ER signature	Non-adjuvant treated patients (n=276)		Patients given endocrine treatment (but not chemotherapy, n=326)	
MCF7 ESR1 siRNA	Not significant		CoxPH p=0.0001 LR p=0.004	
ER+ versus ER-	Not significant		CoxPH p<0.0001 LR p<0.0001	
Simplified tamoxifen predictor	CoxPH p=0.037 LR p=0.194		CoxPH p<0.0001 LR p=0.02	
Endopredict	CoxPH p=0.021 LR p=0.175		CoxPH p<0.0001 LR p<0.0001	
PAM50	CoxPH p=0.154 LR p=0.185		CoxPH p=0.0001 LR p=0.030	

Figure 7 Comparison of ESR1 (Gene Encoding Estrogen Receptor [ER]-α Protein) mRNA Expression, MCF7-ESR1-siRNA (Small Interfering RNA) Principal Component (PC) and ER Status. Log_2 ESR1 mRNA Levels Were Plotted Against MCF7-ESR1-siRNA PC in Breast Tumors Separated by ER Status (Determined by Immunohistochemistry). (A) ER^- Tumors (n = 208) and (B) ER^+ Tumors (n = 822). Patients Who Received Endocrine Therapy Shown in Red and Patients Who Did Not Receive Endocrine Therapy in Black



estrogens produced by the cells. Our PC of estrogen pathway activity was also generated from 1 cell line, MCF7 breast cancer cells. Although MCF7 cells have been used extensively as a model of ER^+ breast cancer and their molecular profiles closely resemble those of luminal breast tumors, the estrogen signaling pathway in MCF7 cells alone might not completely recapitulate estrogen signaling in heterogeneous breast tumors. However, that this approach was successful for the purposes of the present study is supported by the observation that known estrogen-dependent genes were induced or repressed after siRNA reduction of ESR1 in this system.

ESR1 *mRNA Has a Broad Range of Expression Levels in Both ER^+ and ER^- Breast Tumors*

Our analyses showed that ESR1 mRNA is expressed over a range of levels, and this expression overlaps between IHC-determined ER^+ and ER^- tumors, as described in other studies.[18] Also, the range of ESR1 mRNA expression might possibly translate into a range of ER transcriptional activity in a breast tumor that could influence the response to endocrine therapy.[45] The MCF7-ESR1-siRNA first PC correlated with ESR1 mRNA expression levels and was associated with the prognosis of patients treated with endocrine therapies but not that of untreated patients. Luminal A and luminal B tumors have a similar distribution of ESR1 mRNA expression and inferred ER pathway activity according to our PC analysis. This indicates that the different proliferative and clinical behavior of these 2 subtypes is not simply due to differential expression of ESR1 mRNA or differential activity of the ER signaling pathway. Across the other subtypes, a gradation of ESR1 mRNA expression and the magnitude of the MCF7-ESR1-siRNA PC was seen. Compared with ER^+ tumors, most ER^- tumors have much lower ESR1 mRNA expression and reduced ER molecular pathway activity inferred from the magnitude of the MCF7-ESR1-siRNA PC. However, we observed 11 patients whose tumors expressed high ESR1 mRNA levels (ESR1 probe set cutoff ≥ 10; see Materials and Methods) and appeared to have a highly active ER pathway activity (PC cutoff ≥ -0.01; see Materials and Methods) despite being assigned ER^- status using IHC (Figure 3F, red dots).

Studies have supported the use of gene expression profiles to determine ER status, either independently or in concert with ER IHC.[14,19,73] For instance, TargetPrint was developed by Agendia and used to ascertain ESR1, PGR, and ERBB2 levels using gene expression data. TargetPrint had high concordance with the ER and PGR status determined by IHC and ERBB2 status determined by fluorescence in situ hybridization.[74] Bastani et al[17] used a 3-gene classifier (GATA3 [GATA binding protein 3], CA12, and a 701-base pair-expressed sequence tag[17]) to determine ER status using gene expression data. Neither ESR1 nor PGR mRNA expression was used in their classifier, and none of these genes, except for the reference gene, β-actin, are contained in the Onco*type* DX gene list.

However, the criteria for ER positivity in a breast tumor (>1% stained nuclei) represents a local maximum. In contrast, RNA analysis usually represents a global average of a tumor. Therefore, the determination of ER status using gene expression data together with IHC might have particular advantages, which has been proposed by a number of groups,[18,19,73] and seems likely to provide useful additional information for treatment stratification. For example, it was shown that some ER^+ patients with resistance to tamoxifen treatment had low tumor ESR1 mRNA levels.[46] Also, Itoh et al[75] suggested using ESR1 mRNA expression levels to stratify patients with ER^-/PGR^+ breast tumors for endocrine therapy.

Identifying Gene Sets for Use With PC Analysis to Infer ER Pathway Activity

ER-related gene sets have contributed to several prognostic and predictive assays for breast cancer,[46,53,57,76] including the Onco*type*

DX assay and the ECTO (European Cooperative Trial in Operable Breast Cancer) clinical trial.[77] The abundance of single RNA transcripts such as CA12 , NPY1R (neuropeptide Y receptor Y1), STC2 (stanniocalcin 2), DKC1 (dyskerin pseudouridine synthase), or PGR might provide a more readily measurable marker of ER transcriptional activity. However, improved technologies have allowed the analysis of multiple RNAs downstream of a receptor of biologic process to be analyzed accurately and inexpensively, allowing statistical summaries of these RNAs (eg, by PC analysis) to estimate the activity of specific molecular pathways. A key advantage of using a large number of mRNA transcripts to generate a PC is its inherent redundancy and hence robustness in that by having many mRNAs account for a PC, the failure of ≥ 1 measurement can be compensated for by other measured mRNA transcripts. Summarizing multiple tumor RNAs into a single biomarker is increasingly common. For example, a recent study used gene set enrichment analysis to infer the activity of immune gene sets in published breast cancer microarray cohorts.[78]

Our data suggest that the sets of mRNAs differentially expressed in MCF7 cells after siRNA targeting of ESR1, or differentially expressed between tumors identified as ER^+ and ER^- by IHC, are additional candidate gene sets for estimating ER pathway activity in breast tumors. A combination of bioinformatic inference of ER pathway activation using these gene sets, together with information regarding ER-α IHC and ESR1 mRNA expression, might provide a broad-based assessment of breast tumor ER status.

It is interesting that these 2 gene sets and other gene sets previously associated with the estrogen response shared little overlap with one other (Supplemental Figure 4; available in the online version). Presumably this resulted in part from the differing intentions when each of these gene sets was generated (Supplemental Table 3; available in the online version). For example, some of these sets consist of genes summarizing biological pathway activation, which might identify the intactness of the signaling pathway components. Other sets are composed predominantly of genes expressed as a downstream consequence of estrogen signaling and might therefore act as biomarkers for ER activity. Finally, some gene sets were originally designed to be clinically prognostic or predictive. The lack of concordance between ER-associated gene sets does not appear to be solely due to differing levels of statistical stringency. For example, a revised analysis of the mRNAs differentially expressed in MCF7 cells after siRNA targeting of ESR1 using a much less stringently adjusted P value cutoff of $P \leq .05$ did not change the intersection with the small Endopredict gene set (12%; data not shown). For other gene sets, such as the large gene set from Oh et al,[53] relaxing our adjusted P value cutoff to $P \leq .05$ actually reduced the intersection (from 28% to 20%; data not shown). The disparity of the prognostic and predictive gene lists can also be attributed to numerous factors, such as heterogeneity in the patient populations used, the use of different microarray platforms coupled with different probe sets, and different methods of data normalization and sampling variation. Furthermore, this observation could have resulted from the heterogeneity of tumor tissue, in which ESR1-regulated genes are only a small part of the greater part of the breast tumor transcriptome, or the in vitro response of relatively homogenous cultured cell lines might not reproduce estrogen signaling in vivo. Despite their differences, these multiple gene sets might possibly be effective in summarizing different aspects of estrogen pathway activity. At present, IHC hormonal status remains an invaluable tool for breast cancer diagnosis, prognosis, and prediction to endocrine treatment response. This is partly because no reference standard method is available to compare all the different mRNA biomarkers or signatures of estrogen signaling in breast cancer cells. Furthermore, each signature of estrogen activity or signaling that has been identified was designed to capture slightly different aspects of estrogen signaling, rendering it difficult to directly compare and contrast each signature with each other. Our proposed multimodal approach of using IHC ER status, ESR1 mRNA abundance, and a genomic indicator of estrogen pathway activity negates the need to compare 1 type of measurement against another. Prospective clinical studies with larger patient numbers and survival data would provide more information on how these 3 parameters could be used better or weighted, for example, to stratify patients for treatment, and further refinements could be made to improve our multimodal approach by including other relevant molecular or clinical data. With large patient numbers and patient follow-up data, our signatures, along with other published estrogen signatures, could be compared more robustly.

Despite their differences, most of the ER-associated gene sets we studied appeared to have prognostic power. Five estrogen-associated gene sets in particular (MCF7-ESR1-siRNA first PC, genes differentially expressed between ER^+ and ER^- tumors, Endopredict,[55] PAM50,[62] and the simplified tamoxifen predictor[56]) had especially strong prognostic associations across our multicohort data set for patients receiving endocrine therapy, although not for untreated patients (Figure 6), suggesting that they might capture clinically valuable information regarding endocrine treatment response pathways. In addition, the PCs from the MCF7-ESR1-siRNA first PC, the simplified tamoxifen predictor, PAM50, and the Endopredict gene set appeared to add prognostic information to logistic regression analysis over and above the ER IHC status or ESR1 mRNA levels. The abundance of ESR1 mRNA itself had relatively weak associations with patient outcomes in our tumor data set. In future studies, it will be interesting to assess the statistical interactions in prognostic models between ER-associated gene sets and intrinsic subtype, proliferative variables, and other clinicopathologic features.

Breast Tumor Heterogeneity and Utility of Genomic Estimates of ER-α Pathway Activity

Intratumor heterogeneity is a feature of some breast tumors and their metastases,[79] in terms of their cell biology, gene expression, ERBB2 status, and IHC ER status.[80] IHC has a clear advantage over microarray or RT-qPCR analysis for the detection of the ER-α protein in small regions of heterogeneous tumors that might later become clinically important. The proportion of tumor cells to stromal cells in breast tumors varies immensely, and IHC techniques can allow the determination of the percentage of tumor cells in a given section that is assessed. Similarly, the intensity and distribution pattern of the ER-α protein within an already heterogeneous tumor can be accurately detected. In contrast, using mRNA to profile tumors has a limitation, because the heterogeneity in mRNA expression within the cells and the tumor cell versus stromal cell composition within the section is combined and the resulting expression signals

averaged. However, it was shown that despite their heterogeneity, the variation between tumors was greater than the variation within tumors.[81] Furthermore, in addition to inaccuracies in hormone receptor status assignment, possibly due to variations, IHC provides no information regarding the functional status of the ER-α protein or the ER signaling pathway. Multiple other potential mechanisms might also account for the apparent dislocation between ER-α IHC and ER-driven transcriptional responses, such as the action of ER-β and the androgen receptor, and suppression of ER protein cell surface expression despite mRNA expression. Finally, the relationship between the frequency and intensity of ER positivity in tumor cells and treatment response remains unclear. Although some studies have suggested that patients benefited from endocrine therapy, even if as few as 1% of their tumor cells exhibited moderate to strong ER-α protein staining,[15] a recent retrospective study of 9639 patients found that patients with tumors with 1% to 9% staining of ER-α did not appear to benefit as much from endocrine therapy compared with those with ≥ 10% staining.[82] Furthermore, it was also shown that patients with 1% to 9% ER positivity using IHC exhibit a molecular profile similar to that of patients with ER^- tumors, and a high proportion of these tumors are of the basal-like molecular subtype.[14,83] These investigators suggested that the assignment of ER positivity in these tumors might not reflect the biology of the tumors and that these patients might not benefit from endocrine therapy.[83] A variant of ER-α called ERΔ7 might function in a dominant negative manner and is highly expressed in ER^+ basal tumors.[11] Another study identified tumors with 1% to 9% ER positivity by IHC behaving similar to ER^- tumors using RT-qPCR and reported that the survival of patients in this subset of 1% to 9% positivity were intermediate, between that of the remaining ER^+ (>10% positivity) and ER^- tumors.[84]

It has also been argued that with the increased sensitivity of ER antibodies such as SP1 and, hence, the increased bimodal distribution of ER status (90% ER^+), perhaps the range of ER expression is no longer relevant because all suitable patients with > 1% ER positivity will be prescribed endocrine therapy.[16] However, if the assignment of ER status is critical to determine patient response to endocrine therapy and the threshold for ER positivity is > 1% positive staining, it is imperative that the rate of false-negative results during ER status determination using IHC be dramatically reduced.[16] The effect of the established guidelines regarding ER status testing in breast cancer is currently under investigation. However, the effect of these guidelines around ERBB2 status testing in breast cancer has been reviewed. The findings have revealed an overall increase in the ERBB2 positivity rate.[85] Also, a greater proportion of tumors were reassigned to an equivocal ERBB2 status from an $ERBB2^-$ status, which enabled these patients to receive trastuzumab and potentially benefit from this treatment.[86] Thus, it is important to interpret with caution any findings generated from publicly available microarray data sets, because an expectation would exist that a proportion of the tumors would have inaccurate hormone receptor status.

An argument also exists that the bimodal distribution of ER status is not an accurate representation of the true biologic continuum of ER expression.[5,87] It has been reported that tumors with low levels of ER expression (<10% positivity) are difficult to classify,[88,89] with considerable interlaboratory variation for low ER positivity with a false-negative rate of 30% to 60%.[88] Also, a slightly greater discordance was found in tumors with low ER positivity (1%-10% staining) when 2 different ER antibodies and 2 different scoring methods were used.[89] Hence, the determination of ER status, in particular, for tumors with low ER positivity becomes important when stratifying for treatment, especially because the potential benefits of endocrine therapy for this subset of patients are still unclear.

Owing to the heterogeneity and complexity of breast cancer, an integrated and multisystems approach is likely the best method to model this disease, because no single model is likely to recapitulate all aspects of this disease.[90-92] PC analysis in multimodal approaches is valuable when integrating in vitro and in vivo gene expression signatures in breast cancer.[93,94] However, no method has been recommended to integrate in vitro and in vivo data or to integrate clinicopathologic information from conventional tests with molecular-based studies, although a general consensus has been reached that combining both types of information to improve the management of breast cancer, rather than relying on 1 method alone, or advocating the use of prognostic gene expression profiles instead of clinicopathologic prognostic indicators is better.

In our data set, several tumors were IHC ER^+ despite expressing low levels of both ESR1 mRNA and ER-induced genes. These might represent tumors that would benefit from endocrine-based therapy to target ER^+ regions of their heterogeneous areas, such as is current clinical practice. However, we also noted 21 tumors that were IHC ER^- despite apparently having active ER transcriptional pathways, indicated by high levels of both ESR1 mRNA (*ESR1* probe set cutoff, ≥ 6.25; see Materials and Methods) and ER molecular pathway activity (PC cutoff, ≥ −0.01; see Materials and Methods; Figure 7A, black dots). These potentially represent a previously hidden set of patients who have IHC ER^- tumors (hidden possibly because of technical rather than biologic reasons) but might nevertheless benefit from endocrine-based treatment. The presence of ER-α in tumors has been shown to be associated with a good prognosis in patients and also suggests the likelihood of a response to endocrine therapy in ER^+ patients. We suggest that this patient group requires further study. The next sensible step might be to conduct a prospective clinical study with larger numbers of patients and to assess the metastasis-free survival of non—endocrine-treated IHC ER^- patients to determine whether non—endocrine-treated IHC ER^- patients with high ESR1 mRNA expression and/or high estrogen pathway activity PC have better survival outcomes than non—endocrine-treated IHC ER^- patients with lower ESR1 mRNA expression and/or low estrogen pathway activity PC. Similarly, the metastasis-free survival of any ER^+ patients who were not prescribed endocrine therapy could be followed up and assessed to determine whether non—endocrine-treated ER^+ patients with high ESR1 mRNA expression and/or high estrogen pathway activity PC have better survival outcomes than non—endocrine-treated IHC ER^+ patients with lower ESR1 mRNA expression and/or low estrogen pathway activity PC. However, in applying our proposed multimodal method to an ER^+ cohort clinically requires a more careful approach, because it has been shown that patients who are ER^+ have a very good prognosis and those who undergo endocrine therapy benefit from the treatment. Nevertheless, the ER^+ cohort in this context would be clinically useful for investigating the development of endocrine resistance and/or relapse in these patients.

PGR status has been indicative of an intact ER-α signaling pathway in ER^+ patients. However, extrapolating the intactness of the ER-α signaling pathway using IHC PGR alone results in the process of PGR status assignment having the same discrepancies associated with the IHC technique, as discussed previously. The distribution of PGR^+ and PGR^- patients (n = 639) was similar to the distribution of ER^+ and ER^- patients (n = 1034) when the magnitude of our PC of estrogen pathway activity was plotted against *ESR1* mRNA expression (Figure 4F and 5A, respectively). However, similar to our observation of IHC ER^- patients, a proportion of IHC PGR^- patients had high ESR1 mRNA expression and high estrogen pathway activity PC. However, some IHC PGR^- patients also had low PGR mRNA expression but varying magnitudes of estrogen pathway activity PC. It might be possible that some of these IHC PGR^- patients were IHC ER^+; hence, these patients might have been prescribed endocrine therapy. The existence of ER^-/PGR^+ tumors has been questioned, investigated, and confirmed.[95-98] Itoh et al[75] suggested the use of ESR1 mRNA expression to help stratify ER^-/PGR^+ patients, and approximately 25% of these patients' tumors had the molecular profile of ER^+ tumors with similar relapse-free survival to that of ER^+/PGR^+ patients.[75] Shen et al[99] also showed that patients who were ER^-/PGR^+ had a relapse-free survival similar to that of ER^+/PGR^+ patients and had a better response to endocrine therapy than ER^+/PGR^- patients. Multimodal assessments of tumors using a PC of estrogen signaling and ESR1 and/or PGR mRNA expression, together with IHC ER/PGR status, would be extremely useful in stratifying breast cancer patients for endocrine treatment. Schroth et al[97] showed that in ER^-/PGR^+ patients, 59 unique mRNAs were differentially regulated, including ESR1 and GATA3.[97] It is reasonable to assume that PGR status and ER status both correlate with our PC of estrogen pathway activity across our breast tumors. However, further patient stratification is required to investigate these multiple parameters robustly.

It would also be useful to investigate further other clinical aspects of the patients in our data set, such as type of endocrine received (tamoxifen vs. aromatase inhibitors), menopausal status (pre- and postmenopausal), and time to relapse in regard to IHC ER status, ESR1 mRNA abundance, and PC of estrogen signaling. However, the clinical data for the analyzed data sets did not possess this level of detail for us to investigate these factors robustly in our study. In the present study, we built on this work by illustrating the potential complementarity of determining ER status using ESR1 mRNA expression, gene set-based biomarkers for ER pathway activation, and ER IHC in a large and well-annotated data set. We suggest this might be especially helpful to identify patients with IHC ER^- tumors that nevertheless express high levels of both ESR1 mRNA and ER-induced genes. We also suggest that, according to the analyses of multiple ER-associated gene sets (Supplemental Table 3; available in the online version), the information added to ER-α IHC and ESR1 expression by RT-qPCR measurements of single genes such as CA12, NPY1R, STC2, DKC1, and PGR is investigated.

Conclusion

Several gene sets representing ER activity have been previously identified. However, the transcriptional consequences of ER pathway signaling and the biologic meaning of ER gene expression signatures for the prognosis and prediction of endocrine therapy response are still incompletely understood. In addition, previously published gene sets proposed to reflect ER signaling are highly divergent. In the present study, we have shown that despite their differences, most previously published ER-associated gene sets and the 2 new gene sets reported in the present study, have statistically significant associations with patient outcome. The principal components of ER-associated gene lists, or individual highly ER-responsive genes, appear to provide valid indicators of ER pathway activity. We propose that combining the statistical summaries of ER-driven gene sets with ER-α IHC and ESR1 mRNA expression can provide the most robust indicator of ER signaling in the breast tumors of individual patients for research and clinical use.

Clinical Practice Points

- We propose that a 3-way assessment of breast tumor ER biology, combining ER status, ESR1 mRNA expression, and inferred ER pathway activity using PC analysis, could in the future increase the robustness of ER pathway activity estimates in research and when stratifying patients for endocrine-based therapy.
- We suggest that a "hidden" set of patients might exist who have ER^- tumors by IHC but also have active ER transcriptional pathways (indicated by high levels of both ESR1 mRNA and ER-induced genes) and thus might potentially benefit from endocrine-based therapy.

Acknowledgments

We wish to thank Mr Liam Williams from the Centre of Genomics and Proteomics, Auckland, New Zealand for his assistance with microarray hybridization. For the New Zealand breast tumor analysis, we are grateful to the New Zealand Breast Cancer Registers for their provision of patient follow-up data; the surgical teams in Auckland led by Mr Wayne Jones and Dr Vanessa Blair for the collection of patient samples at surgery after informed patient consent; Drs Reena Ramsaroop and Robyn Oldfield for pathological assistance; and Mrs Helen Morrin, Dr Bridget Robinson, and the staff of the Christchurch Tissue Bank for provision of tissue bank breast cancer material and accompanying clinical data. We thank Breast Cancer Cure, New Zealand, for supporting this work.

Disclosure

The authors declare that they have no competing interests.

Supplemental Data

Supplemental data accompanying this article can be found in online version at http://dx.doi.org/10.1016/j.clbc.2016.09.001.

References

1. Jemal A, Bray F, Center MM, et al. Global cancer statistics. *CA Cancer J Clin* 2011; 61:69-90.
2. Manavathi B, Dey O, Gajulapalli VN, et al. Derailed estrogen signaling and breast cancer: an authentic couple. *Endocr Rev* 2013; 34:1-32.
3. De Marchi T, Foekens JA, Umar A, Martens JW. Endocrine therapy resistance in estrogen receptor (ER)-positive breast cancer. *Drug Discov Today* 2016; 21:1181-8.
4. Hammond ME, Hayes DF, Wolff AC, Mangu PB, Temin S. American Society of Clinical Oncology/College of American Pathologists guideline recommendations

for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. *J Oncol Pract* 2010; 6:195-7.
5. Allred DC. Issues and updates: evaluating estrogen receptor-alpha, progesterone receptor, and HER2 in breast cancer. *Mod Pathol* 2010; 23:S52-9.
6. Nadji M. Quantitative immunohistochemistry of estrogen receptor in breast cancer: "much ado about nothing!". *Appl Immunohistochem Mol Morphol* 2008; 16:105-7.
7. Diaz LK, Sneige N. Estrogen receptor analysis for breast cancer: current issues and keys to increasing testing accuracy. *Adv Anat Pathol* 2005; 12:10-9.
8. Sauter G, Lee J, Bartlett JM, Slamon DJ, Press MF. Guidelines for human epidermal growth factor receptor 2 testing: biologic and methodologic considerations. *J Clin Oncol* 2009; 27:1323-33.
9. Wolff AC, Hammond ME, Schwartz JN, et al. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. *Arch Pathol Lab Med* 2007; 131:18-43.
10. Barone I, Brusco L, Fuqua SA. Estrogen receptor mutations and changes in downstream gene expression and signaling. *Clin Cancer Res* 2010; 16:2702-8.
11. Groenendijk FH, Zwart W, Floore A, Akbari S, Bernards R. Estrogen receptor splice variants as a potential source of false-positive estrogen receptor status in breast cancer diagnostics. *Breast Cancer Res Treat* 2013; 140:475-84.
12. Wolff AC, Hammond ME, Hicks DG, et al. Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update. *Arch Pathol Lab Med* 2014; 138:241-56.
13. Hammond ME. ASCO-CAP guidelines for breast predictive factor testing: an update. *Appl Immunohistochem Mol Morphol* 2011; 19:499-500.
14. Iwamoto T, Booser D, Valero V, et al. Estrogen receptor (ER) mRNA and ER-related gene expression in breast cancers that are 1% to 10% ER-positive by immunohistochemistry. *J Clin Oncol* 2012; 30:729-34.
15. Fisher ER, Anderson S, Dean S, et al. Solving the dilemma of the immunohistochemical and other methods used for scoring estrogen receptor and progesterone receptor in patients with invasive breast carcinoma. *Cancer* 2005; 103:164-73.
16. Calhoun BC, Collins LC. Predictive markers in breast cancer: an update on ER and HER2 testing and reporting. *Semin Diagn Pathol* 2015; 32:362-9.
17. Bastani M, Vos L, Asgarian N, et al. A machine learned classifier that uses gene expression data to accurately predict estrogen receptor status. *PLoS One* 2013; 8: e82144.
18. Gong Y, Yan K, Lin F, et al. Determination of oestrogen-receptor status and ERBB2 status of breast carcinoma: a gene-expression profiling study. *Lancet Oncol* 2007; 8:203-11.
19. Roepman P, Horlings HM, Krijgsman O, et al. Microarray-based determination of estrogen receptor, progesterone receptor, and HER2 receptor status in breast cancer. *Clin Cancer Res* 2009; 15:7003-11.
20. Iverson AA, Gillett C, Cane P, et al. A single-tube quantitative assay for mRNA levels of hormonal and growth factor receptors in breast cancer specimens. *J Mol Diagn* 2009; 11:117-30.
21. Bordeaux JM, Cheng H, Welsh AW, et al. Quantitative in situ measurement of estrogen receptor mRNA predicts response to tamoxifen. *PLoS One* 2012; 7: e36559.
22. Badve SS, Baehner FL, Gray RP, et al. Estrogen- and progesterone-receptor status in ECOG 2197: comparison of immunohistochemistry by local and central laboratories and quantitative reverse transcription polymerase chain reaction by central laboratory. *J Clin Oncol* 2008; 26:2473-81.
23. Ellis MJ, Perou CM. The genomic landscape of breast cancer as a therapeutic roadmap. *Cancer Discov* 2013; 3:27-34.
24. Gyorffy B, Hatzis C, Sanft T, et al. Multigene prognostic tests in breast cancer: past, present, future. *Breast Cancer Res* 2015; 17:11.
25. Nik-Zainal S, Davies H, Staaf J, et al. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature* 2016; 534:47-54.
26. Gao Y, Jones A, Fasching PA, et al. The integrative epigenomic-transcriptomic landscape of ER positive breast cancer. *Clin Epigenetics* 2015; 7:126.
27. Morganella S, Alexandrov LB, Glodzik D, et al. The topography of mutational processes in breast cancer genomes. *Nat Commun* 2016; 7:11383.
28. Tishchenko I, Milioli HH, Riveros C, Moscato P. Extensive transcriptomic and genomic analysis provides new insights about luminal breast cancers. *PLoS One* 2016; 11:e0158259.
29. Ali HR, Rueda OM, Chin SF, et al. Genome-driven integrated classification of breast cancer validated in over 7,500 samples. *Genome Biol* 2014; 15:431.
30. Mertins P, Mani DR, Ruggles KV, et al. Proteogenomics connects somatic mutations to signalling in breast cancer. *Nature* 2016; 534:55-62.
31. Pereira B, Chin SF, Rueda OM, et al. The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes. *Nat Commun* 2016; 7:11479.
32. Marcotte R, Sayad A, Brown KR, et al. Functional genomic landscape of human breast cancer drivers, vulnerabilities, and resistance. *Cell* 2016; 164:293-309.
33. Sun Y, Liu Q. Deciphering the correlation between breast tumor samples and cell lines by integrating copy number changes and gene expression profiles. *Biomed Res Int* 2015; 2015:901303.
34. Harbeck N, Sotlar K, Wuerstlein R, Doisneau-Sixou S. Molecular and protein markers for clinical decision making in breast cancer: today and tomorrow. *Cancer Treat Rev* 2014; 40:434-44.
35. Azim HA Jr, Michiels S, Zagouri F, et al. Utility of prognostic genomic tests in breast cancer practice: the IMPAKT 2012 Working Group Consensus Statement. *Ann Oncol* 2013; 24:647-54.
36. Park MM, Ebel JJ, Zhao W, Zynger DL. ER and PR immunohistochemistry and HER2 FISH versus Onco*type* DX: implications for breast cancer treatment. *Breast J* 2014; 20:37-45.
37. Wallden B, Storhoff J, Nielsen T, et al. Development and verification of the PAM50-based Prosigna breast cancer gene signature assay. *BMC Med Genomics* 2015; 8:54.
38. Fitzal F, Filipits M, Rudas M, et al. The genomic expression test EndoPredict is a prognostic tool for identifying risk of local recurrence in postmenopausal endocrine receptor-positive, Her2neu-negative breast cancer patients randomised within the prospective ABCSG 8 trial. *Br J Cancer* 2015; 112:1405-10.
39. Mouttet D, Lae M, Caly M, et al. Estrogen-receptor, progesterone-receptor and HER2 status determination in invasive breast cancer: concordance between immuno-histochemistry and MapQuant microarray based assay. *PLoS One* 2016; 11:e0146474.
40. Wesseling J, Tinterri C, Sapino A, et al. An international study comparing conventional versus mRNA level testing (TargetPrint) for ER, PR, and HER2 status of breast cancer. *Virchows Arch* 2016; 469:297-304.
41. Nguyen B, Cusumano PG, Deck K, et al. Comparison of molecular subtyping with BluePrint, MammaPrint, and TargetPrint to local clinical subtyping in breast cancer patients. *Ann Surg Oncol* 2012; 19:3257-63.
42. Allott EH, Cohen SM, Geradts J, et al. Performance of three-biomarker immunohistochemistry for intrinsic breast cancer subtyping in the AMBER consortium. *Cancer Epidemiol Biomarkers Prev* 2016; 25:470-8.
43. Kok M, Linn SC. Gene expression profiles of the oestrogen receptor in breast cancer. *Neth J Med* 2010; 68:291-302.
44. Poola I, Yue Q. Estrogen receptor alpha (ER alpha) mRNA copy numbers in immunohistochemically ER alpha-positive-, and negative breast cancer tissues. *BMC Cancer* 2007; 7:56.
45. Giulianelli S, Molinolo A, Lanari C. Targeting progesterone receptors in breast cancer. *Vitam Horm* 2013; 93:161-84.
46. Kim C, Tang G, Pogue-Geile KL, et al. Estrogen receptor (ESR1) mRNA expression and benefit from tamoxifen in the treatment and prevention of estrogen receptor-positive breast cancer. *J Clin Oncol* 2011; 29:4160-7.
47. Hatanaka Y, Nagasaki M, Yamaguchi R, et al. A novel strategy to search conserved transcription factor binding sites among coexpressing genes in human. *Genome Inform* 2008; 20:212-21.
48. Ma S, Kosorok MR. Identification of differential gene pathways with principal component analysis. *Bioinformatics* 2009; 25:882-9.
49. Tsai CH, Chiu JH, Yang CW, et al. Molecular characteristics of recurrent triple-negative breast cancer. *Mol Med Rep* 2015; 12:7326-34.
50. Zhang Y, Zhang J, Liu Z, Liu Y, Tuo S. A network-based approach to identify disease-associated gene modules through integrating DNA methylation and gene expression. *Biochem Biophys Res Commun* 2015; 465:437-42.
51. von der Heyde S, Wagner S, Czerny A, et al. mRNA profiling reveals determinants of trastuzumab efficiency in HER2-positive breast cancer. *PLoS One* 2015; 10: e0117818.
52. Thakkar AD, Raj H, Chakrabarti D, et al. Identification of gene expression signature in estrogen receptor positive breast carcinoma. *Biomark Cancer* 2010; 2: 1-15.
53. Oh DS, Troester MA, Usary J, et al. Estrogen-regulated genes predict survival in hormone receptor-positive breast cancers. *J Clin Oncol* 2006; 24:1656-64.
54. Symmans WF, Hatzis C, Sotiriou C, et al. Genomic index of sensitivity to endocrine therapy for breast cancer. *J Clin Oncol* 2010; 28:4111-9.
55. Filipits M, Rudas M, Jakesz R, et al. A new molecular predictor of distant recurrence in ER-positive, HER2-negative breast cancer adds independent information to conventional clinical risk factors. *Clin Cancer Res* 2011; 17:6012-20.
56. Loi S, Haibe-Kains B, Desmedt C, et al. Predicting prognosis using molecular profiling in estrogen receptor-positive breast cancer treated with tamoxifen. *BMC Genomics* 2008; 9:239.
57. Cheng WY, Ou Yang TH, Anastassiou D. Biomolecular events in cancer revealed by attractor metagenes. *PLoS Comput Biol* 2013; 9:e1002920.
58. Lasham A, Samuel W, Cao H, et al. YB-1, the E2F pathway, and regulation of tumor cell growth. *J Natl Cancer Inst* 2012; 104:133-46.
59. Hu Z, Fan C, Oh DS, et al. The molecular portraits of breast tumors are conserved across microarray platforms. *BMC Genomics* 2006; 7:96.
60. Xu X, Veenstra TD. Concentration of endogenous estrogens and estrogen metabolites in the NCI-60 human tumor cell lines. *Genome Med* 2012; 4:31.
61. Berthois Y, Katzenellenbogen JA, Katzenellenbogen BS. Phenol red in tissue culture media is a weak estrogen: implications concerning the study of estrogen-responsive cells in culture. *Proc Natl Acad Sci U S A* 1986; 83:2496-500.
62. Parker JS, Mullins M, Cheang MC, et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. *J Clin Oncol* 2009; 27:1160-7.
63. Nelander S, Wang W, Nilsson B, et al. Models from experiments: combinatorial drug perturbations of cancer cells. *Mol Syst Biol* 2008; 4:216.
64. Tegner J, Bjorkegren J. Perturbations to uncover gene networks. *Trends Genet* 2007; 23:34-41.
65. Bild AH, Yao G, Chang JT, et al. Oncogenic pathway signatures in human cancers as a guide to targeted therapies. *Nature* 2006; 439:353-7.
66. Kao J, Salari K, Bocanegra M, et al. Molecular profiling of breast cancer cell lines defines relevant tumor models and provides a resource for cancer gene discovery. *PLoS One* 2009; 4:e6146.
67. Ross DT, Scherf U, Eisen MB, et al. Systematic variation in gene expression patterns in human cancer cell lines. *Nat Genet* 2000; 24:227-35.
68. Musgrove EA, Caldon CE, Barraclough J, Stone A, Sutherland RL. Cyclin D as a therapeutic target in cancer. *Nat Rev Cancer* 2011; 11:558-72.

69. Teschendorff AE, Gomez S, Arenas A, et al. Improved prognostic classification of breast cancer defined by antagonistic activation patterns of immune response pathway modules. *BMC Cancer* 2010; 10:604.
70. Sas L, Lardon F, Vermeulen PB, et al. The interaction between ER and NFkappaB in resistance to endocrine therapy. *Breast Cancer Res* 2012; 14:212.
71. Gruvberger-Saal SK, Bendahl PO, Saal LH, et al. Estrogen receptor beta expression is associated with tamoxifen response in ERalpha-negative breast carcinoma. *Clin Cancer Res* 2007; 13:1987-94.
72. Ross-Innes CS, Stark R, Teschendorff AE, et al. Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* 2012; 481: 389-93.
73. Li Q, Eklund AC, Juul N, et al. Minimising immunohistochemical false negative ER classification using a complementary 23 gene expression signature of ER status. *PLoS One* 2010; 5:e15031.
74. Viale G, Slaets L, Bogaerts J, et al. High concordance of protein (by IHC), gene (by FISH; HER2 only), and microarray readout (by TargetPrint) of ER, PgR, and HER2: results from the EORTC 10041/BIG 03-04 MINDACT trial. *Ann Oncol* 2014; 25:816-23.
75. Itoh M, Iwamoto T, Matsuoka J, et al. Estrogen receptor (ER) mRNA expression and molecular subtype distribution in ER-negative/progesterone receptor-positive breast cancers. *Breast Cancer Res Treat* 2014; 143:403-9.
76. Schmidt M, Victor A, Bratzel D, et al. Long-term outcome prediction by clinicopathological risk classification algorithms in node-negative breast cancer—comparison between Adjuvant!, St Gallen, and a novel risk algorithm used in the prospective randomized node-negative-breast cancer-3 (NNBC-3) trial. *Ann Oncol* 2009; 20:258-64.
77. Bianchini G. Proliferation-, estrogen-, and T-cell-related metagenes to predict outcome after adjuvant/neoadjuvant chemotherapy for operable breast cancer in the ECTO trial. *J Clin Oncol* 2013; 31, abstract 1014.
78. Chang YH, Chiu YC, Hsu YC, et al. Applying gene set analysis to characterize the activities of immune cells in estrogen receptor positive breast cancer. *Transl Cancer Res* 2016; 5:176-85.
79. Polyak K. Heterogeneity in breast cancer. *J Clin Invest* 2011; 121:3786-8.
80. Beca F, Polyak K. Intratumor heterogeneity in breast cancer. *Adv Exp Med Biol* 2016; 882:169-89.
81. Barry WT, Kernagis DN, Dressman HK, et al. Intratumor heterogeneity and precision of microarray-based predictors of breast cancer biology and clinical outcome. *J Clin Oncol* 2010; 28:2198-206.
82. Yi M, Huo L, Koenig KB, et al. Which threshold for ER positivity? A retrospective study based on 9639 patients. *Ann Oncol* 2014; 25:1004-11.
83. Deyarmin B, Kane JL, Valente AL, et al. Effect of ASCO/CAP guidelines for determining ER status on molecular subtype. *Ann Surg Oncol* 2013; 20:87-93.
84. Prabhu JS, Korlimarla A, Desai K, et al. A majority of low (1-10%) ER positive breast cancers behave like hormone receptor negative tumors. *J Cancer* 2014; 5: 156-65.
85. Varga Z, Noske A. Impact of modified 2013 ASCO/CAP guidelines on HER2 testing in breast cancer: one year experience. *PLoS One* 2015; 10:e0140652.
86. Overcast WB, Zhang J, Zynger DL, Tozbikian GH. Impact of the 2013 ASCO/CAP HER2 revised guidelines on HER2 results in breast core biopsies with invasive breast carcinoma: a retrospective study. *Virchows Arch* 2016; 469: 203-12.
87. Allred DC, Mohsin SK. ER expression is not bimodal in breast cancer. *Am J Clin Pathol* 2005; 124:474-6.
88. Rhodes A, Jasani B, Barnes DM, Bobrow LG, Miller KD. Reliability of immunohistochemical demonstration of oestrogen receptors in routine practice: interlaboratory variance in the sensitivity of detection and evaluation of scoring systems. *J Clin Pathol* 2000; 53:125-30.
89. Reisenbichler ES, Lester SC, Richardson AL, et al. Interobserver concordance in implementing the 2010 ASCO/CAP recommendations for reporting ER in breast carcinomas: a demonstration of the difficulties of consistently reporting low levels of ER expression by manual quantification. *Am J Clin Pathol* 2013; 140:487-94.
90. Vargo-Gogola T, Rosen JM. Modelling breast cancer: one size does not fit all. *Nat Rev Cancer* 2007; 7:659-72.
91. Bild AH, Parker JS, Gustafson AM, et al. An integration of complementary strategies for gene-expression analysis to reveal novel therapeutic opportunities for breast cancer. *Breast Cancer Res* 2009; 11:R55.
92. Chang JT, Carvalho C, Mori S, et al. A genomic strategy to elucidate modules of oncogenic pathway signaling networks. *Mol Cell* 2009; 34:104-14.
93. Starmans MH, Chu KC, Haider S, et al. The prognostic value of temporal in vitro and in vivo derived hypoxia gene-expression signatures in breast cancer. *Radiother Oncol* 2012; 102:436-43.
94. Rees MG, Seashore-Ludlow B, Cheah JH, et al. Correlating chemical sensitivity and basal gene expression reveals mechanism of action. *Nat Chem Biol* 2016; 12: 109-16.
95. De Maeyer L, Van Limbergen E, De Nys K, et al. Does estrogen receptor negative/progesterone receptor positive breast carcinoma exist? *J Clin Oncol* 2008; 26:335-8.
96. Rhodes A, Jasani B. The oestrogen receptor-negative/progesterone receptor-positive breast tumour: a biological entity or a technical artefact? *J Clin Pathol* 2009; 62:95-6.
97. Schroth W, Winter S, Buttner F, et al. Clinical outcome and global gene expression data support the existence of the estrogen receptor-negative/progesterone receptor-positive invasive breast cancer phenotype. *Breast Cancer Res Treat* 2016; 155:85-97.
98. Hefti MM, Hu R, Knoblauch NW, et al. Estrogen receptor negative/progesterone receptor positive breast cancer is not a reproducible subtype. *Breast Cancer Res* 2013; 15:R68.
99. Shen T, Brandwein-Gensler M, Hameed O, Siegal GP, Wei S. Characterization of estrogen receptor-negative/progesterone receptor-positive breast cancer. *Hum Pathol* 2015; 46:1776-84.